PERGAMON Computers and Mathematics with Applications 37 (1999) 87–97

An International Journal
computers & mathematics with applications

Local Convexity Results in a Generalized Fermat-Weber Problem

J. BRIMBERG
School of Business, University of Prince Edward Island
Charlottetown, Prince Edward Island, C1A 4P3, Canada

R. F. LOVE
Department of Management Science/Systems, McMaster University
Hamilton, Ontario L8S 4M4, Canada

(Received and accepted June 1998)

Abstract—A generalized form of the Fermat-Weber problem requires finding a point in $\mathbb{R}^N$ to minimize a sum of nondecreasing functions of distances to m given points. In this paper, local convexity properties are investigated for the generalized problem. Sufficient conditions are derived which guarantee that the Hessian matrix of the objective function will be positive definite. The analysis also reveals that Weiszfeld-type iterative algorithms may have sublinear convergence rates, since the Hessian may only be positive semidefinite at a local minimum.

Keywords—Fermat-Weber problem, Convexity, Convergence.

1. INTRODUCTION

The Fermat-Weber problem requires finding a point in $\mathbb{R}^N$ which minimizes the sum of weighted distances to m given points. In a practical setting, the unknown point represents the location of a new facility, and the given points are customers (also referred to as destinations, existing facilities, demand points, fixed points, or vertices). The weighted distance between the new facility and a customer is a measure of the cost to transport goods or provide services by the facility to satisfy the customer's known requirements. More generally, the cost components may be nonlinear, nondecreasing functions of distance, so that the formulation of our model becomes

$$\text{Minimize}\, W(x) = \sum_{i=1}^{m} g_i(d(x, a_i)), \tag{1}$$

where $a_i = (a_{i1}, \dots, a_{iN})^\top$ is the known position of the i^{th} customer, $i = 1, \dots, m$; m gives the number of customers; $x = (x_1, \dots, x_N)^\top$ is the unknown position of the new facility; $d(x, y)$ is a distance function which measures the distance separating any two points $x, y \in \mathbb{R}^N$; and $g_i(u)$ is a nondecreasing function of the scalar u in the interval $[0, +\infty)$, for $i = 1, \dots, m$.

This research was supported by grants from the Natural Science and Engineering Research Council of Canada and the Academic Research Program of the Department of National Defence of Canada.

0898-1221/99/\$ - see front matter Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX
PII: S0898-1221(99)00102-9

In the standard form, the cost components are given by

$$g_i(d(x,a_i)) = w_i d(x,a_i), \qquad i = 1,\dots,m, \tag{2}$$

where the w_i are positive weighting constants, and $d(x,y)$ is assumed to be a norm (typically the Euclidean norm). Convexity properties of the objective function $W(x)$ are well known for the Fermat-Weber problem in standard form. An early attempt by Haley [1] considers a weighted sum of Euclidean distances in $\mathbb{R}^2$. Convexity of the objective function is shown by an analysis of the second-order derivatives. An alternate proof is given by Love [2] to avoid the difficulty resulting from the nondifferentiability of $W(x)$ at the fixed points a_i. His proof relies on the triangle inequality to show that the Euclidean norm is a convex function of $x \in \mathbb{R}^2$. This result readily generalizes to any norm on $\mathbb{R}^N$.

In a seminal paper on convergence properties in the Fermat-Weber problem, Kuhn [3] notes that the objective function, given by a weighted sum of Euclidean distances in $\mathbb{R}^N$, is strictly convex when the fixed points are not collinear. This result readily follows from the observation that the Euclidean norm ($\ell_2(x) = (x_1^2 + \cdots + x_N^2)^{(1/2)}$) is strictly convex along any straight line not passing through the origin ($x = 0$). The stronger property allows us to conclude that the solution x^* which minimizes $W(x)$ is unique. Strict convexity of the objective function can be extended to a broad class of norms referred to either as round norms [4] or S-norms [5]. It also readily follows that the preceding results hold when $g_i(u)$ is an increasing convex function, $\forall\, i = 1,\dots,m$, (e.g., see [5–7]).

The local convexity of $W(x)$ in its general form may be verified by examining the Hessian matrix,

$$H(x) = \begin{bmatrix} \dfrac{\partial^2 W}{\partial x_1^2} & \cdots & \dfrac{\partial^2 W}{\partial x_N \partial x_1} \\ \vdots & & \vdots \\ \dfrac{\partial^2 W}{\partial x_1 \partial x_N} & \cdots & \dfrac{\partial^2 W}{\partial x_N^2} \end{bmatrix}. \tag{3}$$

If $H(x)$ is positive definite at a given point x, then $W(x)$ is strictly convex in a δ-neighborhood of x; however, the converse may not be true. El-Shaieb [8] examines a (2×2) Hessian matrix for the case $g_i(u) = w_i u^k$, $i = 1,\dots,m$, and $d(x,y)$ is the Euclidean norm on $\mathbb{R}^2$. Positive definiteness is demonstrated for $k \ge 1$, by showing that $\frac{\partial^2 W}{\partial x_1^2}$, $\frac{\partial^2 W}{\partial x_2^2}$ and the determinant of $H(x)$ are all positive-valued functions.

To motivate the general cost structure in (1), we note that a considerable amount of research has taken place recently on empirically-based distance-predicting functions. Previously, the majority of researchers in the field of location theory believed that actual travel distances could be accurately predicted by either the rectangular (ℓ_1) or Euclidean (ℓ_2) norm (see for example, [9]). However, two sets of empirical studies by Love and Morris [10, 11] showed that considerable gains in accuracy can be made by using the weighted ℓ_p norm rather than either the weighted ℓ_1 or ℓ_2 norms. This result is not surprising since the ℓ_1 and ℓ_2 norms are special cases of the ℓ_p norm.

The development of the concept of rectangular bias in road networks by Brimberg and Love [12] led to a further improvement in the predictive power of the weighted ℓ_p norm. It was shown in empirical studies by Brimberg, Love and Walker [13] that a significant improvement in the accuracy of the weighted ℓ_p norm can be achieved by rotating the axes until they are aligned with the rectangular bias in the road network of the geographical region under consideration.

In the present paper, we consider the travel cost of a journey to be a nondecreasing twice-differentiable function of the ℓ_p norm. The empirical underpinnings of this assumption arise from several sources. Love and Morris [10, 11] propose a distance function which is the ℓ_p norm raised to a power. Their empirical studies showed that this function (with an extra parameter) is superior to the ℓ_p norm. The authors generally observe economies of scale in the road networks examined, although a few cases show diseconomies. These results make intuitive sense. One would usually expect an economy of scale arising from the larger number of routes available when points are further apart.

The i^{th} term in the summation given in equation (1) is generally regarded as the shipping cost from the facility to be located to the i^{th} customer. In the empirical study by Westwood [14], it is shown that the cost of making a shipment can be broken down into a variable cost per mile together with a fixed charge component which is the prorated sum of items such as weekly driver cost, monthly vehicle cost, and other annual fixed transport costs. Such a concave cost function commonly occurs in manufacturing and logistical operations (see also [15]).

Kolesar, Walker and Hausner [16] have developed a model for determining the travel time of a fire engine in response to calls as follows:

$$T(d) = \begin{cases} 2\left(\dfrac{d}{a}\right)^{1/2}, & \text{if } d \leq 2d_c, \\ \dfrac{v_c}{a} + \dfrac{d}{v_c}, & \text{if } d > 2d_c, \end{cases}$$

where the travel time is T; the travel distance between the fire station and the calling point is d; the acceleration is a; the cruising velocity is v_c; and the distance required to achieve cruising velocity is d_c. Once the point coordinates of the calling point are determined, a prediction of the travel distance can be calculated by using a distance predicting function. Thus, an estimated travel time, which takes into account the road network of the region, is determined. This type of travel time model, or a variation of it, can be used for other types of vehicles, such as ambulances, school buses, and police cars. We note that $T(d)$ is a concave function of d.

With the preceding motivation in mind, the objective of this paper will be to examine the properties of the Hessian matrix $H(x)$ under the more general cost structure found in (1). The analysis extends the work of El-Shaieb [8] as follows:

(i) the location problem occurs in $\mathbb{R}^N$ (i.e., $H(x)$ is an $(N \times N)$ matrix);
(ii) $g_i(u)$ is any nondecreasing, twice-differentiable function of $u \geq 0$, for $i = 1, \dots, m$;
(iii) the distance function $d(x, y)$ can be any ℓ_p norm.

General results are obtained for positive-definiteness of $H(x)$. These results are useful for generalizing convexity properties of the objective function, and as shown here, for determining local convergence rates of iterative algorithms such as the Weiszfeld procedure and its extensions (e.g., see [7, 17–19]).

2. PRELIMINARY RESULTS

We consider the following problem:

$$\text{Minimize } W(x) = \sum_{i=1}^{m} g_i\left(\ell_p(x - a_i)\right), \tag{4}$$

where the distance function is now given by the ℓ_p norm. More precisely, the distance between any two points $x, a_i \in \mathbb{R}^N$ is a function of the coordinates of the vector $(x - a_i)$ as follows:

$$\ell_p(x - a_i) = \left[\sum_{t=1}^{N} |x_t - a_{it}|^p\right]^{1/p}, \qquad p \geq 1. \tag{5}$$

The restriction on the parameter p ensures that $\ell_p(\cdot)$ is a norm. Note that when $p = 1$, we have the rectangular norm (also known as Manhattan or city-block distance), and when $p = 2$, the Euclidean or straight-line distance.

The first- and second-order partial derivatives of g_i with respect to the coordinates x_t are determined using the chain rule of calculus.

$$\frac{\partial}{\partial x_t} g_i\left(\ell_p(x)\right) = g_i'\left(\ell_p(x)\right)\frac{\partial \ell_p(x)}{\partial x_t}, \qquad t = 1, \ldots, N, \tag{6}$$

$$\frac{\partial^2}{\partial x_j \partial x_t} g_i\left(\ell_p(x)\right) = g_i'\left(\ell_p(x)\right)\frac{\partial^2 \ell_p(x)}{\partial x_j \partial x_t} + g_i''\left(\ell_p(x)\right)\frac{\partial \ell_p(x)}{\partial x_j}\frac{\partial \ell_p(x)}{\partial x_t}, \tag{7}$$

$$\forall j, t = 1, \ldots, N, \quad i = 1, \ldots, m,$$

where $g_i'(u)$ and $g_i''(u)$ denote, respectively, the first- and second-order derivatives of $g_i(u)$, which are assumed to exist $\forall u \geq 0$. The partial derivatives of $\ell_p(x)$ are readily obtained as follows:

$$\frac{\partial \ell_p(x)}{\partial x_t} = \frac{\operatorname{sign}(x_t)|x_t|^{p-1}}{[\ell_p(x)]^{p-1}}, \qquad t = 1, \ldots, N, \tag{8}$$

$$\frac{\partial^2 \ell_p(x)}{\partial x_t^2} = \frac{(p-1)\,|x_t|^{p-2}\sum_{j\neq t}|x_j|^p}{[\ell_p(x)]^{2p-1}}, \qquad t = 1, \ldots, N, \tag{9}$$

and

$$\frac{\partial^2 \ell_p(x)}{\partial x_j \partial x_t} = \frac{-(p-1)\operatorname{sign}(x_j)\operatorname{sign}(x_t)\,|x_j|^{p-1}\,|x_t|^{p-1}}{[\ell_p(x)]^{2p-1}}, \tag{10}$$

$$\forall j, t, \qquad j \neq t.$$

Properties 1 and 2 below are immediately evident from relations (8), (9), and (10).

PROPERTY 1. *If $p \geq 1$, the first-order partial derivatives of $\ell_p(x)$ are defined everywhere except at the origin ($x = 0$). If $p = 1$, they are defined everywhere except on the union of hyperplanes given by*

$$S = \bigcup_{t=1}^{N} \{x \mid x_t = 0\}. \tag{11}$$

PROPERTY 2. *If $p \geq 2$, the second-order partial derivatives of $\ell_p(x)$ are defined $\forall x \neq 0$. If $1 \leq p < 2$, they are defined $\forall x \in \mathbb{R}^N \setminus S$.*

Referring to (6) and (7), it is clear that the preceding results apply equally well to the first- and second-order partial derivatives of the functions $g_i(\ell_p(x))$.

We now consider the Hessian matrix of $g_i(\ell_p(x))$, denoted by $H_i(x)$ for $i = 1, \ldots, m$. The elements of H_i are given by

$$h_{tj}^{(i)}(x) = h_{jt}^{(i)}(x) = \frac{\partial^2}{\partial x_j \partial x_t}\left[g_i\left(\ell_p(x)\right)\right], \quad \forall j, t. \tag{12}$$

Let $y = (y_1, \ldots, y_N)^\top$ denote any nonzero vector in $\mathbb{R}^N$, and consider the quadratic form $Q_i(x; y) = y^\top H_i(x) y$. At any point x where the partial derivatives in (12) exist, we obtain

$$\begin{aligned} Q_i &= \sum_{t=1}^{N}\sum_{j=1}^{N} h_{jt}^{(i)}(x) y_j y_t \\ &= Q_{i1} + Q_{i2}, \end{aligned} \tag{13}$$

where

$$Q_{i1} = g_i'(\ell_p(x))A(x;y), \tag{14}$$

$$Q_{i2} = g_i''(\ell_p(x))B(x;y), \tag{15}$$

$$A(x;y) = \sum_{t=1}^{N}\sum_{j=1}^{N} \frac{\partial^2 \ell_p(x)}{\partial x_j \partial x_t} y_j y_t, \tag{16}$$

and

$$B(x;y) = \left[\sum_{t=1}^{N} \frac{\partial \ell_p(x)}{\partial x_t} y_t\right]^2. \tag{17}$$

To analyse the quadratic form in more detail, the following three cases need to be investigated.

(i) $1 \le p < 2$.

Here all the elements of x are nonzero in order for the Hessian matrix to be defined. Hence, we can write

$$y_t = b_t x_t, \quad t = 1, \dots, N, \tag{18}$$

where the b_t are scalar quantities, not all of which are zero. Substituting (9), (10), and (18) into (16) gives

$$\begin{aligned} A(x;y) &= \sum_{t=1}^{N} (p-1) \left| \frac{|x_t|^{p-2} \sum_{j\neq t} |x_j|^p}{[\ell_p(x)]^{2p-1}} \right| y_t^2 - \sum\sum_{j\neq t} \frac{(p-1)\operatorname{sign}(x_t)\operatorname{sign}(x_j)\,|x_t|^{p-1}\,|x_j|^{p-1}}{[\ell_p(x)]^{2p-1}} y_t y_j \\ &= (p-1)\sum\sum_{j\neq t} \frac{|x_t|^p\,|x_j|^p\,(b_t^2 - b_t b_j)}{[\ell_p(x)]^{2p-1}} \\ &= \frac{(p-1)}{[\ell_p(x)]^{2p-1}} \sum\sum_{j<t} |x_t|^p\,|x_j|^p\,(b_t - b_j)^2. \end{aligned} \tag{19}$$

(ii) $p = 2$.

The requirement on x is relaxed, so that some (but not all) of the x_t may now be zero. Letting $y_t = b_t x_t$, $\forall\, x_t \neq 0$, we obtain

$$\begin{aligned} A(x;y) &= \sum\sum_{j\neq t} \frac{x_j^2 y_t^2}{[\ell_p(x)]^3} - \sum\sum_{j\neq t} \frac{x_t x_j y_t y_j}{[\ell_p(x)]^3} \\ &= \frac{1}{[\ell_p(x)]^3} \left(\sum\sum_{\substack{j<t \\ x_j, x_t \neq 0}} x_t^2 x_j^2 (b_t - b_j)^2 + \sum\sum_{\substack{j\neq t \\ x_t = 0}} x_j^2 y_t^2 \right). \end{aligned} \tag{20}$$

(iii) $p > 2$.

In similar fashion,

$$A(x;y) = \frac{(p-1)}{[\ell_p(x)]^{2p-1}} \sum\sum_{\substack{j<t \\ x_j, x_t \neq 0}} |x_t|^p\,|x_j|^p\,(b_t - b_j)^2. \tag{21}$$

Recalling that $p \ge 1$, we conclude from (19), (20), and (21) that $A(x;y) \ge 0$. From (17), it is seen that $B(x;y) \ge 0$. Furthermore, since $g_i(u)$ is a nondecreasing function of u, $g_i'(u) \ge 0$, so that $Q_{i1}(x;y) \ge 0$. The sign of $Q_{i2}(x;y)$ will depend on the sign of the second derivative $g_i''(u)$. Thus, depending on the form of the cost components g_i, $i = 1, \dots, m$, the following general results are obtained.

PROPERTY 3. *If $g_i(u)$ is a convex function $\forall u \geq 0$, then $Q_i(x;y) \geq 0$, $\forall y$ and $x \in \mathbb{R}^N$ where $H_i(x)$ is defined.*

PROOF. Since $g_i(u)$ is convex $\forall u \geq 0$, it follows that $g_i''(\ell_p(x)) \geq 0, \forall x$. Therefore, $Q_{i2}(x;y) \geq 0$, and $Q_i(x;y)$ is the sum of two nonnegative terms.

PROPERTY 4. *Let $g_i(u)$ be a concave function $\forall u \geq 0$, and consider any point x where $H_i(x)$ is defined. In general, $Q_i(x;y)$ will be $>$, $=$, or < 0 depending on the direction of y.*

PROOF. Since $g_i(u)$ is concave, it follows that $g_i''(\ell_p(x)) \leq 0$, $\forall x$. Thus, $Q_{i2}(x;y) \leq 0$, $\forall y$. We can choose y as a scalar multiple of x ($y = bx$), in which case $A(x;y) = 0$, and $Q_i(x;y) = Q_{i2}(x;y) \leq 0$. Alternatively, let y be tangent to the contour of $\ell_p(x)$ at $x(\nabla \ell_p(x) \cdot y = 0$, where ∇ denotes the gradient vector), so that $B(x;y) = 0$, and $Q_i(x;y) = Q_{i1}(x;y) \geq 0$. Since $Q_i(x;y)$ varies continuously with y, we conclude that $Q_i(x;y)$ is indeterminate in sign.

Property 3 implies that $H_i(x)$ is positive semidefinite when $g_i(u)$ is convex, from which we infer that $g_i(\ell_p(x))$ is convex with respect to x. This result could be stated directly, since it is well known that a nondecreasing convex function of a convex function is also convex. From Property 4, it follows that $H_i(x)$ is indefinite when $g_i(u)$ is concave, and hence, $g_i(\ell_p(x))$ is neither convex nor concave at x. The next result provides sufficient conditions for $Q_i(x;y)$ to be strictly positive, which will be required in the subsequent analysis.

PROPERTY 5. *Let $g_i(u)$ be an increasing convex function with a strictly positive first-order derivative $\forall u > 0$. Then $Q_i(x;y) > 0$ if any one of the following sets of conditions is satisfied:*

a) *$p = 1$, $g_i''(\ell_p(x)) > 0$, and $\nabla \ell_p(x) \cdot y \neq 0$;*
b) *$1 < p \leq 2$, and either $g_i''(\ell_p(x)) > 0$ or $y \neq bx$ for any scalar b;*
c) *$p > 2$, the set $A = \{t \mid x_t \neq 0\}$ has cardinality ≥ 2, and either $g_i''(\ell_p(x)) > 0$ or a scalar b does not exist such that $y_t = bx_t$, $\forall t \in A$.*

PROOF. We have $g_i'(u) > 0$ and $g_i''(u) \geq 0$, $\forall u > 0$. If the conditions in (a) apply, it is clear that $Q_i(x;y) = Q_{i2}(x;y) > 0$. If $1 < p \leq 2$, then by (19) and (20), $A(x;y) = 0$ if and only if $b_t = b$, $\forall t$. However, if $y = bx$, then y cannot be orthogonal to $\nabla \ell_p(x)$, and we conclude that $A(x;y)$ and $B(x;y)$ cannot both be zero simultaneously. Hence, $Q_i(x;y) > 0$ if (b) applies. Similar reasoning holds for (c).

3. MAIN RESULTS

Let $H(x)$ denote the Hessian matrix of the objective function $W(x)$ given in (4). It follows that

$$H(x) = \sum_{i=1}^{m} H_i(x - a_i), \tag{22}$$

or alternatively, the elements of $H(x)$ are given by

$$h_{jt}(x) = \sum_{i=1}^{m} h_{jt}^{(i)}(x - a_i), \quad \forall j, \quad t = 1, \dots, N. \tag{23}$$

Consider first the singular points where $H(x)$ is undefined. Letting

$$S_i = \bigcup_{t=1}^{N} \{x \mid x_t - a_{it} = 0\}, \quad i = 1, \dots, m, \tag{24}$$

and

$$\Omega = \begin{cases} \bigcup_{i=1}^{m} S_i, & \text{if } 1 \leq p < 2, \\ \{a_1, \dots, a_m\}, & \text{if } p \geq 2, \end{cases} \tag{25}$$

we obtain immediately from Property 2 the following result.

PROPERTY 6. *$H(x)$ is defined everywhere except at the set of singular points given by Ω.*

For the remainder of the analysis, it will be assumed that only nonsingular points of the Hessian matrix are being considered. We will investigate the quadratic form, $Q(x;y) = y^\top H(x)y$, where once again, y denotes any vector other than the zero vector.

Using (22),

$$Q(x;y) = \sum_{i=1}^{m} y^\top H_i(x-a_i)\, y = \sum_{i=1}^{m} Q_i(x-a_i;y)\,. \tag{26}$$

General results can now be derived from Property 5 which provide sufficient conditions for $Q(x;y) > 0$, $\forall\, y$, and hence, for $H(x)$ to be positive definite.

PROPERTY 7. *Let $g_i(u)$ be an increasing, convex function with a strictly positive first-order derivative $\forall\, u > 0$, $i = 1,\dots,m$, and let $p \in (1,2]$. If the fixed points $a_1,\dots,a_m$ are not collinear, then $Q(x;y) > 0$, $\forall\, y \neq 0$, $x \in \mathbb{R}^N \setminus \Omega$.*

PROOF. From Property 3, it follows that $Q_i(x-a_i;y) \geq 0$, $\forall\, i$, so that $Q(x;y)$ is the sum of m nonnegative terms. It suffices to show that at least one of these terms is strictly positive. If $Q_i(x-a_i;y) = 0$ for some $i \in \{1,\dots,m\}$, we conclude from Property 5 that a nonzero scalar b can be found such that $y = b(x-a_i)$. However, since the fixed points are not all collinear, there must be at least one index k such that $(x-a_k)$ is not a scalar multiple of $(x-a_i)$. Hence, $Q(x;y) \geq Q_k(x-a_k;y) > 0$.

In the standard form of the Fermat-Weber problem, the cost components are linear functions of distance; i.e., $g_i(u) = w_i u$, $\forall\, i = 1,\dots,m$, where the w_i are positive weights. We see from Property 7 that $H(x)$ is positive definite when the fixed points are not collinear, $\forall\, x \in \mathbb{R}^N \setminus \Omega$. This implies an alternate proof of the well-known result that $W(x)$ is strictly convex in this case. In addition, the positive definiteness of the Hessian matrix is a key factor in determining the local convergence rates of iterative solution procedures such as the Weiszfeld algorithm (e.g., [19]).

PROPERTY 8. *The noncollinearity condition on the set of fixed points in Property 7 is not required if $g_i''(u) > 0$, $\forall\, u > 0$, for at least one $i \in \{1,\dots,m\}$.*

PROOF. From Property 5, it follows that $Q_i(x-a_i;y)$ is strictly positive at all times. Hence, $Q(x;y) \geq Q_i(x-a_i;y) > 0$, $\forall\, y \neq 0$, $x \in \mathbb{R}^N \setminus \Omega$. ∎

When $p > 1$, the ℓ_p norm belongs to a class referred to as round norms which are characterized by contours containing no flat spots. For $p = 1$ (rectangular distances), we have a block norm which is characterized by polytope contours. (For a comparison of these two classes of norms, see [20].) For this reason, the rectangular norm must be treated as a special case. The next result follows immediately from Property 5.

PROPERTY 9. *Let $p = 1$, and consider any $x \in \mathbb{R}^N \setminus \Omega$. A necessary condition for $H(x)$ to be positive definite is that $\{\nabla \ell_p(x-a_i),\ i = 1,\dots,m\}$ contains a basis of $\mathbb{R}^N$. If $g_i''(u) > 0$ in the interval $(0,+\infty)$, $\forall\, i = 1,\dots,m$, then $H(x)$ will be positive definite $\forall\, x \in \mathbb{R}^N \setminus \Omega$ where the preceding condition is satisfied.*

A simple geometric interpretation can be applied to the preceding result. For example, in two-dimensional space ($\mathbb{R}^2$), draw a horizontal and vertical line through the point x parallel to the reference axes, and label the four resulting quadrants counterclockwise as numbers I, II, III, and IV. Then, the set of vectors, $\nabla \ell_1(x-a_i)$, $i = 1,\dots,m$, will form a basis in $\mathbb{R}^2$, if, and only if, the fixed points $a_1,\dots,a_m$ are not all contained within diagonally-opposite quadrants I and III, or II and IV. Equivalently, we see that if the fixed points $a_1,\dots,a_m$ are all contained within a pair of diagonally-opposite quadrants (say I and III), then $W(x)$ will be constant on a 45° line segment through the other pair of diagonally-opposite quadrants (II and IV); hence, $W(x)$ cannot be strictly convex at x.

Results which are analogous to Properties 7 and 8 can be derived for the case $p > 2$. However, there is one anomaly to be noted. Suppose that two lines can be found parallel to a pair of reference axes, which intersect all the fixed points $a_1, \dots, a_m$, and intersect each other at some point B. It is well known that $W(x)$ is strictly convex for any $p > 1$, whenever all the g_i are convex functions and the a_i are noncollinear. However, when $p > 2$, the Hessian matrix $H(B)$ may only be positive semidefinite since $A(B - a_i; y) = 0$, $\forall y$ and $i = 1, \dots, m$.

As a simple illustration, consider a problem in $\mathbb{R}^2$ with four fixed points given by $a_1 = (-4,0)^\top$, $a_2 = (4,0)^\top$, $a_3 = (0,-1)^\top$, and $a_4 = (0,1)^\top$. Let the objective function be in standard form, with $g_i(u) = w_i u$, $i = 1, \dots, 4$, and $w_1 = w_2$, $w_3 = w_4$. Then $W(x)$ is a strictly convex function of x with a unique optimal solution at $x^* = (0,0)^\top$. However, $Q(x^*; y) = 0$, $\forall y$. This demonstrates a basic result which may not often be appreciated, namely that strict convexity of a function within a region does not imply its Hessian matrix is positive definite at every point in that region. It is also interesting to note that a sublinear convergence rate was observed when the Weiszfeld algorithm was used to solve this example. Sublinear rates have been noted previously when x^* coincides with a fixed point [18,19], but not when x^* is at a differentiable point of the objective function as above.

When the cost components g_i are not all convex functions, the quadratic form $Q(x; y)$ will be greater than, equal, or less than zero in general depending on x and the direction y. This result follows from Property 4. Alternatively, we can say that $H(x)$ is indefinite. It also follows that the Hessian matrix may only be positive semidefinite at a local minimum of the objective function. As discussed in the next section, this in turn implies a sublinear convergence rate for Weiszfeld-type iterative algorithms.

4. APPLICATION TO THE WEISZFELD ITERATIVE PROCEDURE

Proceeding in the same manner as for the standard model (e.g., see [19]), we consider the necessary conditions for a stationary, differentiable point of the objective function. Setting the first-order partial derivatives of the objective function in (4) to zero, the following one-point iterative scheme is readily obtained:

$$x_t^{q+1} = f_t(x^q), \qquad t = 1, \dots, N, \tag{27}$$

where

$$f_t(x) = \frac{\sum_{i=1}^{m} v_i(x)\, |x_t - a_{it}|^{p-2}\, a_{it} / [\ell_p(x - a_i)]^{p-1}}{\sum_{i=1}^{m} v_i(x)\, |x_t - a_{it}|^{p-2} / [\ell_p(x - a_i)]^{p-1}}, \qquad t = 1, \dots, N, \tag{28}$$

$$v_i(x) = g_i'\left(\ell_p(x - a_i)\right), \qquad i = 1, \dots, m, \tag{29}$$

and the superscript $q = 0, 1, 2, \dots,$ denotes the iteration number.

The iteration functions f_t have the same form as in the standard model, except that $v_i(x) = w_i$, $\forall i$, in the standard model, whereas now the v_i are functions of x given by (29). Without loss of generality, let us assume that the g_i are all strictly increasing functions, so that

$$v_i(x) > 0, \qquad \forall x \in \mathbb{R}^N, \quad i = 1, \dots, m. \tag{30}$$

Using a similar procedure as in [21] for the standard model, it readily follows that the mapping $F(x) : x \to (f_1(x), \dots, f_N(x))^\top$ is compact. Furthermore, for any value of the parameter p in the closed interval $[1, 2]$, the descent property holds; i.e., $W(x^{q+1}) < W(x^q)$, if $x^{q+1} \neq x^q$. We may conclude as a result that an infinite sequence of distinct iterates generated by F will converge to a local minimum of the objective function for any $p \in [1, 2]$.

In order to investigate the rate of convergence in the neighbourhood of a local minimum $x^* = (x_1^*, \dots, x_N^*)^\top$, which is not a singular point of the iteration functions, we need to consider the Jacobian matrix of first-order derivatives given by

$$F'(x^*) = \begin{bmatrix} \dfrac{\partial f_1(x^*)}{\partial x_1} & \dots & \dfrac{\partial f_1(x^*)}{\partial x_N} \\ \vdots & & \vdots \\ \dfrac{\partial f_N(x^*)}{\partial x_1} & \dots & \dfrac{\partial f_N(x^*)}{\partial x_N} \end{bmatrix}. \tag{31}$$

For convenience, the following notation is used:

$$y_{it}(x) = \frac{v_i(x)\,|x_t - a_{it}|^{p-2}}{[\ell_p(x - a_i)]^{p-1}}, \qquad \forall\, i, t, \tag{32}$$

$$s_t(x) = \sum_{i=1}^{m} y_{it}(x), \qquad \forall\, t. \tag{33}$$

Using a similar derivation as in [19] for the standard model, it is readily shown that

$$\frac{\partial f_t(x^*)}{\partial x_j} = \frac{1}{s_t(x^*)} \sum_{i=1}^{m} \left[\frac{\partial y_{it}(x^*)}{\partial x_j}\right] \cdot (a_{it} - x_t^*), \qquad \forall\, j, t. \tag{34}$$

Observing that

$$\begin{aligned} y_{it}(x) \cdot (a_{it} - x_t) &= \frac{-v_i(x)\,\mathrm{sign}\,(x_t - a_{it})\,|x_t - a_{it}|^{p-1}}{[\ell_p(x - a_i)]^{p-1}} \\ &= -\frac{\partial}{\partial x_t} g_i\left(\ell_p(x - a_i)\right), \end{aligned}$$

it follows that

$$\left[\frac{\partial}{\partial x_j} y_{it}(x)\right] \cdot (a_{it} - x_t) = \begin{cases} -\dfrac{\partial^2}{\partial x_j \partial x_t} g_i\left(\ell_p(x - a_i)\right), & \text{if } j \neq t, \\ \dfrac{\partial^2}{\partial x_t^2} g_i\left(\ell_p(x - a_i)\right) + y_{it}(x), & \text{if } j = t. \end{cases} \tag{35}$$

Combining (34) and (35), we arrive at the following fundamental relation:

$$F'(x^*) = I - S(x^*)H(x^*), \tag{36}$$

where I is the identity matrix, $H(x)$ is the Hessian matrix from the preceding section, and $S(x)$ is a diagonal matrix with positive diagonal elements given by

$$S(x) = \begin{bmatrix} (s_1(x))^{-1} & 0 & \dots & 0 \\ 0 & (s_2(x))^{-1} & \dots & 0 \\ \vdots & & & \\ 0 & \dots & & (s_N(x))^{-1} \end{bmatrix}. \tag{37}$$

If $H(x^*)$ is not positive definite, at least one of its eigenvalues must be zero. Hence, we can show that at least one of the eigenvalues of $F'(x^*)$ equals unity, which implies that convergence to x^* will occur at sublinear rates. In order for local convergence rates to be linear, $H(x^*)$ must be strictly positive definite.

As an illustration of sublinear convergence rates, we consider the following problem in $\mathbb{R}^2$. There are four fixed points at the corners of a square,

$$a_1 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^{\mathsf{T}}, \qquad a_2 = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^{\mathsf{T}},$$
$$a_3 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^{\mathsf{T}}, \quad \text{and} \quad a_4 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^{\mathsf{T}},$$

with associated cost components given by

$$g_i(u) = 1 + (u-1)^3, \qquad u \geq 0, \quad i = 1, \dots, 4,$$

where u denotes the distance to the new facility. Note that $g_i(u)$ is concave for $0 \leq u \leq 1$, and convex for $u \geq 1$, and that $g_i'(1) = g_i''(1) = 0$. This type of cost function might model a situation where there are economies of scale up to a certain distance, beyond which an alternate mode of transportation would be required with rapidly increasing costs.

Using the Euclidean norm to measure distances, the objective function for the above problem takes the form,

$$W(x) = 4 + \sum_{i=1}^{4} \left[\ell_2(x - a_i) - 1\right]^3.$$

It is readily verified that $x^* = (0,0)^{\mathsf{T}}$ minimizes W, and that the elements of $H(x^*)$ are all equal to zero (signifying that W is very flat in the neighbourhood of x^*). Adapting the Weiszfeld procedure (27) to solve the problem gives the following set of recursive relations:

$$x_t^{q+1} = \frac{\sum_{i=1}^{4} u_i(x^q)\, a_{it}}{\sum_{i=1}^{4} u_i(x^q)}, \qquad q = 0, 1, 2, \dots,$$

where

$$u_i(x) = \frac{\left[\ell_2(x - a_i) - 1\right]^2}{\ell_2(x - a_i)}, \qquad i = 1, \dots, 4,$$

and x^0 is an arbitrarily chosen starting point. As expected, testing of this iterative scheme revealed sublinear convergence rates to x^*.

5. CONCLUSIONS

Convexity properties of a generalized form of the Fermat-Weber problem are investigated in terms of the Hessian matrix of the objective function. Sufficient conditions are derived which guarantee that the Hessian will be positive definite. This provides an alternate proof of well-known convexity results for the Fermat-Weber problem. The analysis also shows that the Hessian matrix may only be positive semidefinite at a local or global minimum. This implies that sublinear convergence rates are possible even at nonsingular points of the iteration functions, for Weiszfeld-type iterative solution procedures.

REFERENCES

1. K.B. Haley, The siting of depots, *The International Journal of Production Research* **2** (1), 41–45, (1963).
2. R.F. Love, A note on the convexity of the problem of siting depots, *The International Journal of Production Research* **2** (2), 153–154, (1967).
3. H.W. Kuhn, A note on Fermat's problem, *Mathematical Programming* **4**, 98–107, (1973).
4. J. Brimberg and R.F. Love, Generalized hull properties for location problems, *IIE Transactions* **27**, 226–232, (1995).

5. B. Pelegrin, C. Michelot and F. Plastria, On the uniqueness of optimal solutions in continuous location theory, *European Journal of Operational Research* **20**, 327–331, (1985).
6. R.F. Love and J.G. Morris, On convexity proofs in location theory, *Naval Research Logistics Quarterly* **25** (1), 179–181, (1978).
7. J.G. Morris, Convergence of the Weiszfeld algorithm for Weber problems using a generalized 'distance' function, *Operations Research* **29** (1), 37–48, (1981).
8. A.M. El-Shaieb, The single source Weber problem—Survey and extensions, *Journal of the Operational Research Society* **29** (5), 469–476, (1978).
9. R.L. Francis, Some aspects of a minimax location problem, *Operations Research* **15** (6), 1163–1169, (1967).
10. R.F. Love and J.G. Morris, Modelling inter-city road distances by mathematical models, *Operational Research Quarterly* **23** (1), 61–71, (1972).
11. R.F. Love and J.G. Morris, Mathematical models of road travel distances, *Management Science* **25** (2), 130–139, (1979).
12. J. Brimberg and R.F. Love, General considerations on the use of the weighted ℓ_p norm as an empirical distance measure, *Transportation Science* **27** (4), 341–349, (1993).
13. J. Brimberg, R.F. Love and J.H. Walker, The effect of axis rotation on distance estimation, *European Journal of Operational Research* **80**, 357–364, (1995).
14. J.B. Westwood, A transport planning model for primary distribution, *Interfaces* **8** (1), 1–10, (1977).
15. M.A. Efroymson and T.L. Ray, A branch and bound algorithm for plant location, *Operations Research* **14** (3), 361–368, (1966).
16. P. Kolesar, W. Walker and J. Hausner, Determining the relation between fire engine travel times and travel distances in New York City, *Operations Research* **23** (4), 614–627, (1975).
17. I.N. Katz, On the convergence of a numerical scheme for solving some locational equilibrium problems, *SIAM Journal of Applied Mathematics* **17** (6), 1224–1231, (1969).
18. I.N. Katz, Local convergence in Fermat's problem, *Mathematical Programming* **6**, 89–104, (1974).
19. J. Brimberg and R.F. Love, Local convergence in a generalized Fermat-Weber problem, *Annals of Operations Research* **40**, 33–66, (1992).
20. J.-F. Thisse, J.E. Ward and R.E. Wendell, Some properties of location problems with block and round norms, *Operations Research* **32** (6), 1309–1327, (1984).
21. J. Brimberg and R.F. Love, Global convergence of a generalized iterative procedure for the minisum location problem with ℓ_p distances, *Operations Research* **41** (11), 1153–1163, (1993).